DOCUMENT RESUME

ED 046 973 TM 000 333

AUTHOR McNeil, Keith A.; Beggs, Donald L.

TITLE Directional Hypotheses with the Multiple Linear Regression Approach.

PUB DATE Feb 71

NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29

DESCRIPTORS Correlation, *Hypothesis Testing, *Multiple Regression Analysis, *Research Methodology, *Statistical Analysis, *Tests of Significance

ABSTRACT

Two well known directional (one-tailed) tests of significance, mean difference and correlation coefficient, are presented within the multiple linear regression framework. Adjustments on the computed probability level are indicated. The case for a directional interaction research hypothesis is defended. Conservative adjustments on the computed probability level are offered and a more precise computation is requested of statisticians. Emphasis is placed more on the research question being asked than on blind adherence to conventional formulae. (Author)

Directional Hypotheses With the Multiple Linear Regression Approach

Keith A. McNeil and Donald L. Beggs
Southern Illinois University

Abstract

Two well known directional tests of significance are presented within the multiple linear regression framework. Adjustments on the computed probability level are indicated. The case for a directional interaction research hypothesis is defended. Conservative adjustments on the computed probability level are offered and a more precise computation is requested of statisticians. Emphasis is placed more on the research question being asked than on blind adherence to conventional formulae.

Introduction

The generalized F ratio within the context of multiple linear regression is known to be applicable to a large number of research questions. There is a class of questions, though, which requires an adjustment in the probability level which is reported by canned computer programs. This reported probability level is for an equally divided "two-tailed" test of significance, but often the researcher has justified a "one-tailed" test of significance. Indeed, whenever the research hypothesis contains directionality, then the required test of significance is "one-tailed." A good deal of the research hypotheses that appear in the literature develop a valid rationale for directionality but very few of them proceed to fully take advantage of their stated alpha level. One only needs to look at, for example, Volume 11 of the Journal of Personality and Social Psychology. Numerous articles in this issue propose directional hypotheses and proceed to use a non-directional test. Indeed, Levinger and Schneider (1969) indicate that the results for one hypothseis was significant in the direction opposite to that hypothesized. In reliability and validity research, the research hypothesis of necessity must be directional. It is seldon that a researcher gets excited about a negative reliability coefficient. Likewise, the researcher hypothesizes the sign of the correlational value indicating validity. A negative correlation would only be expected when two scales are measuring the same phenomenon, but one scale has been reversed. (In this case we would still have all of the critical region in one tail of the sampling distribution.)

There are at least three situations that might require a "one-tailed" test of significance: (1) a research hypothesis suggesting one treatment resulting in a higher mean than another treatment; (2) a research hypothesis specifying either a positive correlation between two variables or a negative correlation between two variables; and (3) a research hypothesis specifying a directional interaction. The first two situations are well documented in the statistical literature, but the last is not mentioned.

Case 1: Directional mean difference research hypothesis

We must be careful in interpreting the probability associated with directional hypotheses because the full and restricted regression models are the same with a one-tailed test as with a two-tailed test. A non-directional research hypothesis would take the form: There is a difference in the mean effect of treatments T_1 and T_2. A directional research hypothesis would take the form: Treatment T_1 results in a larger mean effect than does treatment T_2. The full model in both cases would be:

Model 1: $Y_1 = a_oU + a_1T_1 + a_2T_2 + E_1$; the full model where:

Y_1 = the criterion vector.

U = the unit vector.

T_1 = a 1 if the Y_1 score comes from a person in treatment 1, 0 otherwise.

T_2 = a 1 if the Y_1 score comes from a person in treatment 2, 0 otherwise.

TM 000 333

a_o, a_1, a_2 are weighting coefficients which will produce the smallest sum of squared components in the E_1 vector.

E_1 is the error in prediction, or $(Y_1 - \tilde{Y}_1)$, using the weighting coefficients and the predictor variables in the full model.

For each of the above research hypotheses, the statistical hypothesis is: There is no difference in the (population) treatment means. The statistical hypothesis implies the restriction: $a_1 = a_2$. Forcing this restriction on the full model, we arrive at:

Model 2: $Y_1 = a_oU + E_2$; the restricted model.

All symbols are as defined before, with E_2 being the error in prediction using the weighting coefficients and predictor variables in the restricted model.

The two models can of course be compared with the F test, and the associated probability value will be reported by most canned programs. The probability value is the probability of this large a discrepancy or one larger occuring under the restriction that the two population means are equal. The first two rows in Table 1 indicate the state of affairs when the research hypothesis is non-directional. The reported probability value is for a non-directional test of significance and thus no correction is necessary.

If we are concerned about differences in a given direction, then we must look at the sample means to see if the difference between the means is in the hypothesized direction. If the means are in the direction hypothesized, the third example in Table 1, then we must halve the reported probability level, for it indicates to the researcher how often he would expect this large a discrepancy in both directions. If the means are not in the hypothesized direction (the last example in Table 1), then we surely do not want to hold as tenable the research hypothesis. The correct probability level in this case is $(1 - \frac{PROB}{2})$, where PROB is the reported probability value. Since PROB can never be larger than 1, the smallest actual probability level can never be less than .50, i.e. can never lead to holding as tenable the research hypothesis.

Pedagogically, one might want to illustrate the F distribution as in Figure 1. The top half of the F distribution can be thought of as the F ratios resulting when Treatment 2 has a higher mean than Treatment 1. The bottom half then represents those F ratios resulting when Treatment 1 has a higher mean than Treatment 2. It should be quite clear from Figure 1 that if one's alpha level is .05 the appropriate lower limit for a non-directional test is F = 4.20, whereas if the research hypothesis involves directionality, then F = 2.89 is the appropriate lower limit (this being the tabled F value for alpha = 2 x .05, or for an alpha of .10; degrees of freedom equal 1 and 28).

Case 2: Directional correlational research hypothesis

The argument for this case is similar to the previous argument, the only difference is that here we have a continuous predictor variable rather than a dichotomous predictor variable. Often in correlational research, the research hypothesis is something like: There is a non-zero relationship between X_1 and Y_2. The statistical hypothesis in this case would be: There is a zero relationship between X_1 and Y_2. The full and restricted models would be:

Model 3: $Y_2 = a_oU + a_1X_1 + E_3$; the full model where:

Y_2 = the criterion vector.

U = the unit vector.

X_1 = the continuous predictor vector.

a_o and a_1 are weighting coefficients which will provide the sum of squared component in the E_3 vector.

E_3 is the error in prediction ($Y_2 - \widehat{Y_2}$) using the weighting coefficients and predictor variables in the full model.

The restriction: $a_1 = 0$ results in

Model 4: $Y_2 = a_0U + E_4$; the restricted model where all symbols are as above, and where E_4 is the error in prediction $(Y_2 - \tilde{Y}_2)$ using only the overall mean (a_0).

One's research hypothesis might involve a directional relationship such as: There is a positive correlation between X_1 and Y_2. The full and restricted models would be the same, but again one would have to inspect the sign of the weighting coefficient to make sure the non-zero correlation is in the hypothesized direction. The same kinds of corrections in the probability level are called for in this case as in the case for directional differences, and examples are depicted in Table 2. Indeed, we would expect this to be the case because the test for the difference between two means is algebraically equivalent to the test of significance for the point biserial correlation, a special case of the Pearson Product Moment Correlation (Kelly, Beggs, McNeil, Eichelberger, and Lyon, 1969).

Case 3: Directional interaction research hypothesis

This third case has probably not been utilized in the literature because it has not been described in the standard statistical texts. We are not aware of any applied examples of this case, although many research hypotheses in the literature actually call for such an analysis. When a two-tailed interaction analysis is run on a directional interaction hypothesis rather than the legitimate one-tailed analysis, the researcher is reporting a probability level which is not indicative of the actual pobability. As in the previous cases, if the results are in the hypothesized direction the actual probability value is less than that which the researcher reports. We are not aware of the actual correction, as will be indicated shortly.

An example from the literature may help clarify the problem. Gentile (1968) hypothesized: "the lower the sociocultural level of the student, the more he should benefit from the definition treatment (as compared to the no-definition treatment)." Figure 2 illustrates the kind of interaction indicated by the research hypothesis. Figure 3 illustrates the other half of the situations wherein an interaction can occur. These kinds of interaction in Figure 3 are evidently not of interest to Gentile. Therefore, the reported probability level should be at least halved if the results are in the hypothesized direction.

We say at least halved because there are other kinds of interactions similar to that depicted in Figure 1 which would not reflect the research hypothesis. Figure 4 contains one such situation wherein the definition treatment is inferior to the no-definition treatment. Again one would not want to hold as tenable the research hypothesis with this set of data.

As in the first two cases, the full and restricted models for the directional and non-dirctional interaction questions are exactly the same (See Table 3). The sociocultural levels can be treated as categorical variables or as continuous, and we prefer the latter. (The discussion would become more involved if we didn't do it this way.)

The full model which allows interaction to occur would be:

Model 5: $Y_3 = a_0U + a_1T_1 + a_2T_2 + b_1X_1 + b_2X_2 + E_5$

Where:

Y_3 = the criterion vector.

U = the unit vector.

T_1 = 1 if the subject received the definition treatment, otherwise 0.

T_2 = 1 if the subject received the no-definition treatment, 0 otherwise.

X_1 = sociocultral level of the subject if he received the definition treatment, 0 otherwise.

X_2 = sociocultural level of the subject if he received the no-definition treatment, 0 otherwise.

a_0, a_1, a_2, b_1, b_2 are weighting coefficients which will produce the smallest sum of squared components in the E_5 vector.

E_5 = the error in prediction, $(Y_3 - \tilde{Y}_3)$, using the wighting coefficients and predictor variables in the fuul model.

In this example b_1 and b_2 are the slopes of the straight lines of best fit for the two treatments. The hypothesis of no interaction in the population stipulates that the population slopes are equal ($B_1 = B_2$). Since the sample slopes are the best estimators of the population slopes, the restriction which does not allow interaction to occur is: $b_1 = b_2$. This restriction placed on the full model results in the following restricted model:

Model 6: $Y_3 = a_0U + a_1T_1 + a_2T_2 + b_3X_3 + E_6$

All symbols are as defined above, and where X_3 is the sociocultural level of the subject, no matter which treatment he received. E_6 is the error in prediction, $(Y_3 - \tilde{Y}_3)$, using the weighting coefficients and predictor variables in the restricted model. Again, the full and restricted models can be compared via the generalized F ratio.

If one has a non-directional interaction question and the F is significant then the results can simply be plotted and the reported probability level reported.

If one has a directional interaction question and the F is significant, then the results must be plotted to see if the interaction occurs in the direction hypothesized. If the results are opposite to that hypothesized, we surely would not want to hold as tenable the research hypothesis. If the interaction is in the direction hypothesized, then the exact probability is at least one-half the reported probability.

We feel that the above adjustment is not an exact adjustment, but at this time we are not able to describe the exact probability. We would encourage researchers to consider this question and in the future try to develop the exact probability. Certainly though, the interaction plot must reflect the research hypothesis before the researcher can reject the statistical hypothesis and hold as tenable the research hypothesis.

What we question is the probability statement associated with the interaction test of statistical significance when the researcher has stated a directional interaction research question. The reader should be reminded that the statistical hypothesis when testing either interaction or directional interaction is: There is no interaction, or, the lines are parallel. There are many ways of obtaining interaction and only a small subset of these is of interest to the researcher who is interested in a directional interaction question.

These thoughts seem to be important because many decisions are based on statistical grounds which are being used incorrectly. Many research hypotheses involve a directional hypothesis. The researcher is hurting himself when he uses a two-tailed test rather than a one-tailed test. If his results are in the hypothesized direction, the statistic may not fall in the critical region of the two-tailed test, whereas it might have fallen in the critical region of the one-tailed test. (Please remember to also report the amount of variance being accounted for in either case, as that index will probably communicate more than will the probability value.) What is even more disheartening is to see a researcher develop a beautiful directional hypothesis and then report that his data indicate significance in the "opposite direction." He has used a two-tailed test of significance for the directional hypothesis and has found that the statistic falls in the critical region. A little thought would indicate that the researcher cannot hold as tenable his directional hypothesis under these conditions.

He should report what he found and urge future researchers to develop directional hypotheses to correspond with his data; it is ironical to report something as being significant which was completely opposite to that which was expected. In essence, the rationale behind the directional hypothesis may be incorrect, but that cannot be determined on the initial data.

Table 1

Several hypothetical examples for the
differences between two groups
(Full model takes the form of Model 1 and restricted model the form of Model 2.)

Research Hypothesis	Sample Index	Statistical Hypothesis	Restriction	Sample (a_0+a_1)	Means (a_0+a_2)	Outputted Probability	Correction Needed	Actual Probabil
$\mu_1 \neq \mu_2$	$a_1 \neq a_2$	$\mu_1 = \mu_2$	$a_1 = a_2$	20	15	.07	no correction	.07
$\mu_1 \neq \mu_2$	$a_1 \neq a_2$	$\mu_1 = \mu_2$	$a_1 = a_2$	15	20	.07	no correction	.07
$\mu_1 > \mu_2$	$a_1 \; a_2$	$\mu_1 = \mu_2$	$a_1 = a_2$	20	15	.07	$\frac{PROB}{2}$	.035
$\mu_1 > \mu_2$	$a_1 \; a_2$	$\mu_1 = \mu_2$	$a_1 = a_2$	15	20	.07	$1 - \frac{PROB}{2}$	.965

Table 2

Several hypothetical examples for
correlational hypotheses
(Full Model takes the form of Model 3 and restricted Model the form of Model 4)

Research Hypothesis	Sample Index	Statistical Hypothesis	Restriction	Sample Correlation	Outputted Probability	Correction Needed	Actual Probabilit
$\rho \neq 0$	$a_1 \neq 0$	$\rho = 0$	$a_1 = 0$	.36	.07	no correction	.07
$\rho \neq 0$	$a_1 \neq 0$	$\rho = 0$	$a_1 = 0$	-.36	.07	no correction	.07
$\rho > 0$	$a_1 > 0$	$\rho = 0$	$a_1 = 0$	.36	.07	$\frac{PROB}{2}$	.035
$\rho > 0$	$a_1 > 0$	$\rho = 0$	$a_1 = 0$	-.36	.07	$1 - \frac{PROB}{2}$	.965
$\rho < 0$	$a_1 < 0$	$\rho = 0$	$a_1 = 0$	.36	.07	$1 - \frac{PROB}{2}$	.965
$\rho < 0$	$a_1 < 0$	$\rho = 0$	$a_1 = 0$	-.36	.07	$\frac{PROB}{2}$	.035

Table 3

Several hypothetical examples
for interaction hypotheses
(Full model takes the form of Model 5 and
restricted model the form of Model 6.)

Research Hypothesis	Sample Index	Statistical Hypothesis	Restriction	Sample Values b_1	b_2	Outputted Probability	Correction Needed	Actual Probabi
$\beta_1 \neq \beta_2$	$b_1 \neq b_2$	$\beta_1 = \beta_2$	$b_1 = b_2$	.4	.2	.07	no correction	.07
$\beta_1 \neq \beta_2$	$b_1 \neq b_2$	$\beta_1 = \beta_2$	$b_1 = b_2$	.2	.4	.07	no correction	.07
$\beta_1 > \beta_2$	$b_1 > b_2$	$\beta_1 = \beta_2$	$b_1 = b_2$	.4	.2	.07	$\frac{PROB}{2}$	.035
$\beta_1 > \beta_2$	$b_1 > b_2$	$\beta_1 = \beta_2$	$b_1 = b_2$	.2	.4	.07	$1 - \frac{PROB}{2}$	.965

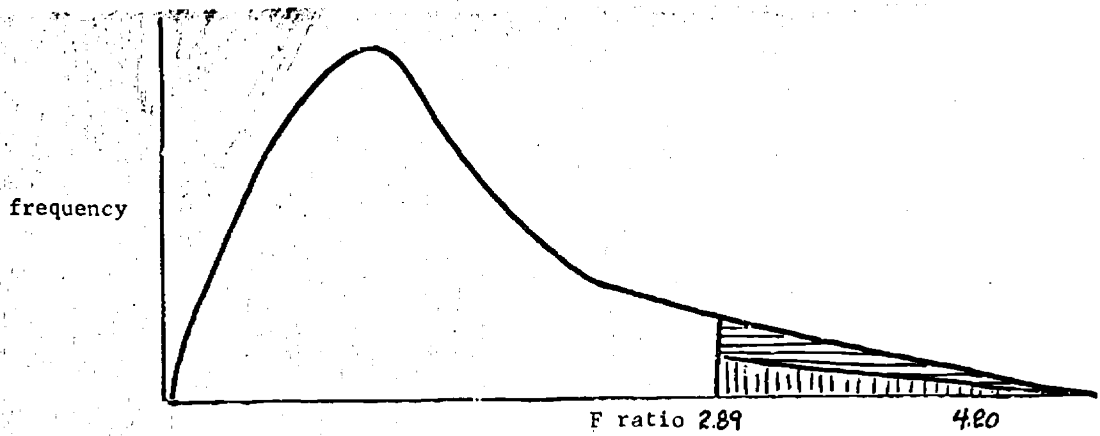

Figure 1

Exemplary F destribution (df_1= 1, df_2 = 28) indicating F ratios resulting under the statistical hypothesis of equal population means. The area depicted by vertical lines represents those F ratios resulting when, say, Treatment 1 has a higher sample mean than Treatment 2. The area depicted by the horizontal lines represents those F ratios resulting when research hypothesis is directional, then the researcher must use the tabled F value for (2 x alpha). This process is analogous to adjusting the reported probability values as indicated in Tables 1, 2, and 3.

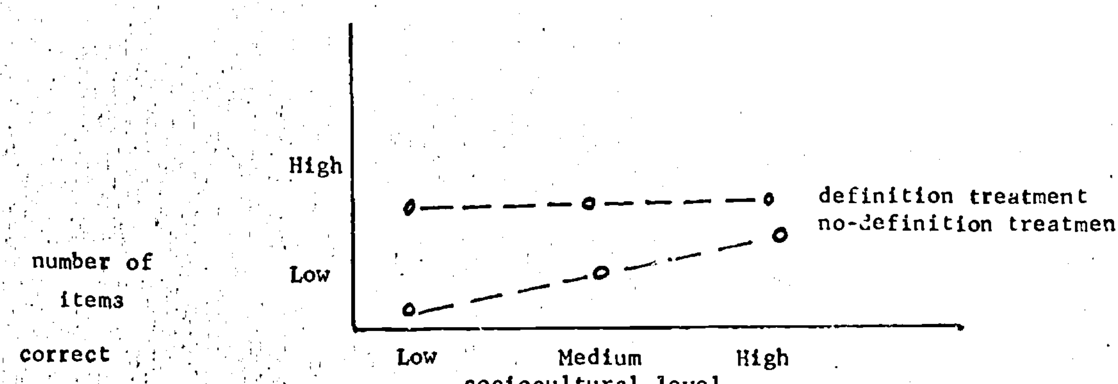

Figure 2

Schematic diagram representing directional interaction hypothesis of Gentile (1968).

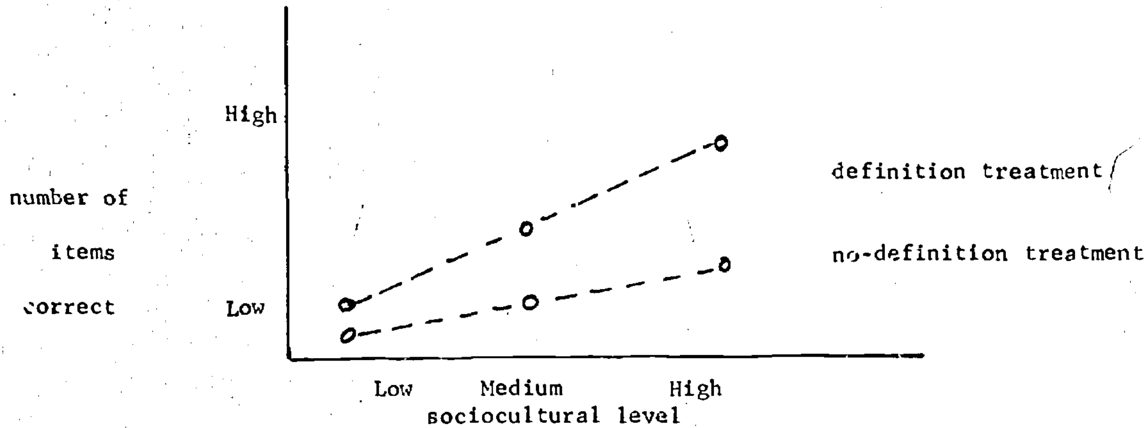

Figure 3

Schematic diagram representing other interactions which could occur but were of no interest to Gentile (1968).

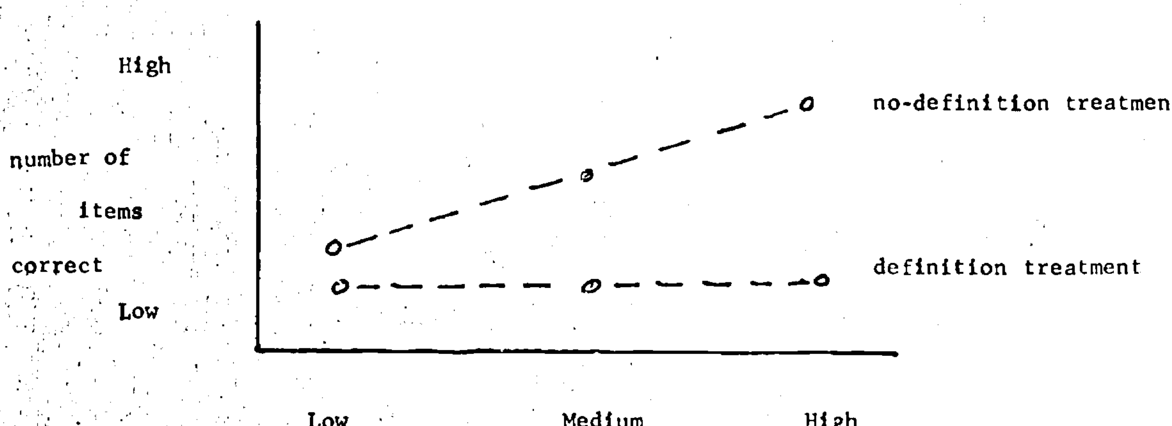

Figure 4

Schematic diagram representing lines similar to Figure 2 but with the definition treatment consistently inferior to the no-definition treatment.

REFERENCES

Gentile, J.R., Sociocultural Level and Knowledge of Definitions in the Solution of Analogy Items, American Educational Research Journal, Vol.5,4, 626-638,1968.

Kelly, F.J., Beggs, D.L., McNeil, K.A., Eichelberger, T., and Lyon, J., Research Design in the Behavioral Sciences: Multiple Regression Approach. Carbondale, Illinois: Southern Illinois University Press, 1969.

Levinger, G. and Schneider, D. J., Test of the "Risk Is a Value" Hypothesis, Journal of Personality and Social Psychology, Vol.II, 2, 165-169, 1969.